UNIVERZA V LJUBLJANI

FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Tendai Evelyn Borin

Nadzor parametrov kakovosti VoIP telefonije

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM PRVE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

Ljubljana, 2017

UNIVERZA V LJUBLJANI

FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Tendai Evelyn Borin

Nadzor parametrov kakovosti VoIP telefonije

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM PRVE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: dr. Andrej Brodnik

Ljubljana, 2017

Rezultati diplomskega dela so intelektualna lastnina avtorja. Za objavljanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Tematika naloge:

Dandanes je dejstvo, da se tudi klasične storitve kot je na primer telefonija tehnološko rešuje z uporabo IP tehnologije. Konkretno VoIP (voice over IP), ki je nadomestek klasične telefonije, je implementiran kot kombinacija protokolov SIP, RTP in RTCP ter še drugih podpornih protokolov. Podjetja, ki nudijo to storitev, morajo zagotavljati visoko stopnjo kakovosti in v ta namen morajo namestiti sisteme za nadzor in upravljanje. V diplomski nalogi preglejte delovanje VoIP storitve s kazalniki njene kakovosti. Poleg tega predlagajte ter implementirajte učinkovito orodje z grafičnim vmesnikom, ki omogoča nadzor prej omenjenih kazalnikov v časovni osi.

Za vodenje in pomoč pri izdelavi diplomskega dela se zahvaljujem mentorju, dr. Andreju Brodniku. Iskrena hvala družini za podporo v času šolanja. Posebno se zahvaljujem tebi, Marko, ki si mi stal ob strani. Zahvaljujem se tudi vsem, ki ste mi pomagali tekom študija.

Kazalo

Seznam uporabljenih kratic

kratica	angleško	slovensko
ACD	Average Call Duration	povprečna dolžina klica
AMQP	Advanced Message Queuing Protocol	protokol za prenos sporočil
AWS	Amazon Web Services	spletna storitev Amazon
CDR	Call Detail Record	razčlenjeni zapis klica
HTTP	Hypertext Transfer Protokol	prenos informacij na spletu
IP	Internet Protocol	internetni protokol
IRC	Internet Relay Chat	spletni klepet
ITU	International telecommunication Union	mednarodna telekomunikacijska zveza
PSTN	Public Switched Telephone Network	javno telefonsko omrežje
QoS	Quality of Service	kakovost storitve
REST	Representational State Transfer	arhitekturni stil spletnih aplikacij
RTT	Round Trip delay Time	čas prenosa
SIP	Session Initiation Protocol	protokol za zagon seje
SMS	Short Sessage Service	sistem kratkih sporočil
SMTP	Simple Mail Transfer Protokol	protokol za prenos elektronske pošte
TCP	Transmission Control Protocol	protokol za nadzor prenosa
UDP	User Datagram Protocol	protokol za prenos paketov
UUID	Universally Unique IDentifier	univerzalen edinstven identifikator
VoIP	Voice over Internet Protocol	govor preko IP

Povzetek

Naslov: Nadzor parametrov kakovosti VoIP telefonije

Centri za nadzor telefonije pri operaterjih in manjša podjetja, ki so posredniki telefonskega prometa, uporabljajo sisteme za spremljanje klicev v realnem času. Pozornost namenjajo kakovosti storitve, ki v veliki meri zavzema nadzor nad kakovostjo večjih vzorcev klicev po različnih omrežjih končnih operaterjev. Cilj diplomskega dela je bil implementirati izboljšavo za nadzor kakovosti VoIP telefonije, ki bi omogočala grafično predstavitev kazalnikov kakovosti klicev. Podani so rezultati raziskave obstoječih orodij na trgu ter definirane zahteve sistema. V diplomski nalogi je predstavljena rešitev z implementacijo sklada ELK za urejanje, statistično obdelavo in grafično predstavitev kazalnikov kakovosti klicev. Pri ovrednotenju rešitve je nakazana možnost nadgradnje s pošiljanjem filtriranih podatkov iz orodja Logstash preko izhoda neposredno na Zabbix strežnik. Podane so ugotovitve, ki izhajajo iz vnaprej postavljenih zahtev sistema.

Ključne besede: VoIP, kazalniki kakovosti, Logstash, Elasticsearch, Kibana, Docker, JSON

Abstract

Title: Monitoring of quality indicators in VoIP

Control centres of telephony operators and other smaller companies that route telephone traffic have a system of regularly monitoring calls in real time. Dedicating a lot of attention to the quality of service, mostly monitoring and controlling quality of larger patterns of calls going to different destinations. The aim of the thesis was to implement improvements in the field of controlling quality of VoIP telephony, enabling visual representations of formal quality indicators of calls such as ASR, NER, ACD. Research was made of the existing tools on the market at the moment and summarized the requirements of the system. The thesis presents a solution of implementing ELK stack for editing data, making statistical analysis and visualization of quality indicators of calls. While evaluating this solution an upgrade is suggested for using Zabbix output of data filtered in Logstash to send them directly to Zabbix server. Findings are given resulting from pre-established system requirements.

Keywords: VoIP, quality parameters, elastic stack, Docker, JSON

Poglavje 1 Uvod

Telekomunikacijska podjetja, ki se ukvarjajo s področjem internetne telefonije, zagotavljajo svojim strankam določeno kakovost klicev. Na spletnih straneh operaterjev in posrednikov telefonskega prometa je mogoče zaslediti, da zagotavljajo glasovne storitve na najvišji ravni kakovosti. To dosegajo s pomočjo sistemov za detekcijo anomalij na večjih vzorcih klicev ter z analizo drugih komunikacijskih storitev.

V diplomskem delu bomo govorili o VoIP telefoniji in kazalnikih kakovosti internetne telefonije. Vsak VoIP klic ima namreč svoje lastnosti, katere lahko uporabimo za statistično obdelavo podatkov. Ker je ročno težko najti korelacije med podatki v tabeli, izvajamo analize in spremljamo spremembe na izbranih parametrih, ki nam pomagajo pri nadzoru kakovosti klicev.

Motivacija za diplomsko nalogo je delo v podjetju Mobik, ki se ukvarja z mednarodnimi govornimi in podatkovnimi storitvami, posluje le s poslovnimi subjekti. Predvsem zagotavlja govorne in podatkovne storitve v mobilnih in fiksnih omrežjih. Omogoča uspešno zaključevanje mednarodnih govornih klicev in sodeluje z največjimi telekomunikacijskimi operaterji. Poleg tega zagotavlja tudi tranzit klicev med operaterji. Razvit ima svoj lasten sistem obračuna in nadzora klicev [25].

1.1 Predstavitev problema

Trenutna predstavitev podatkov o klicih je le v tabelarični obliki, kot je razvidno na sliki 1.1. Za lažje ročno opazovanje ter zaznavo sprememb in morebitnih napak, ki vplivajo na padec kakovosti klicev, bi potrebovali grafično predstavitev podatkov, še posebej kazalnikov kakovosti.

	StartTime	Direction	DNO	OAD	DAD	Duration	Route_id	alerting_start_time	CAUFrom
2	13.1.2017 0:00	O		iiai38163	ii381643	0		0000-00-00 00:00:00	User busy.
3	13.1.2017 0:01	O		iiai38163	ii381643	0		0000-00-00 00:00:00	User busy.
4	13.1.2017 0:02	O		iiai38163	ii381650	24940		13.1.2017 0:02	Normal call clearing.
5	13.1.2017 0:02	O		iiai41794	ii381655	0		13.1.2017 0:02	SIP 487: Request terminated.
6	13.1.2017 0:03	O		iiai22827	ii381644	0		13.1.2017 0:03	Interworking, unspecified.
7	13.1.2017 0:03	O		iiai68663	ii381654	0		13.1.2017 0:03	Normal, unspecified.
8	13.1.2017 0:04	O		iiai22827	ii381644	0		0000-00-00 00:00:00	Interworking, unspecified.
9	13.1.2017 0:05	O		iiai22827	ii381644	0		13.1.2017 0:05	Interworking, unspecified.
10	13.1.2017 0:05	O		iiai38163	ii381643	0		0000-00-00 00:00:00	Normal call clearing.
11	13.1.2017 0:05	O		iiai22827	ii381644	0		0000-00-00 00:00:00	Interworking, unspecified.
12	13.1.2017 0:05	O		iiai59220	ii381654	0		13.1.2017 0:06	Normal, unspecified.
13	13.1.2017 0:06	O		iiai47318	ii381654	0		13.1.2017 0:06	Normal, unspecified.
14	13.1.2017 0:07	O		iiai22827	ii381644	0		0000-00-00 00:00:00	Interworking, unspecified.
15	13.1.2017 0:09	O		iiai15574	ii381654	0		13.1.2017 0:10	Normal, unspecified.
16	13.1.2017 0:13	O		iiai38124	ii381650	8301		0000-00-00 00:00:00	Normal call clearing.
17	13.1.2017 0:18	O		iiai69086	ii381654	0		13.1.2017 0:19	Normal, unspecified.
18	13.1.2017 0:23	O		iiai97237	ii381640	0		13.1.2017 0:23	Normal call clearing.
19	13.1.2017 0:23	O		iiai59226	ii381654	0		13.1.2017 0:23	Normal, unspecified.
20	13.1.2017 0:26	O		iiai59215	ii381654	0		13.1.2017 0:27	Normal, unspecified.
21	13.1.2017 0:28	O		iiai59210	ii381654	0		13.1.2017 0:28	Normal, unspecified.
22	13.1.2017 0:30	O		iiai69059	ii381654	0		13.1.2017 0:30	Normal, unspecified.
23	13.1.2017 0:37	O		iiai97237	ii381640	0		13.1.2017 0:37	Normal call clearing.
24	13.1.2017 0:37	O		iiai12272	ii381644	0		0000-00-00 00:00:00	Interworking, unspecified.
25	13.1.2017 0:38	O		iiai35864	ii381654	0		13.1.2017 0:38	Normal, unspecified.

Slika 1.1: Tabelarična predstavitev podatkov v obstoječem sistemu.

V diplomskem delu predlagamo izboljšavo za nadzor nad kazalniki kakovosti klicev. Rešitev bomo implementirali s skladom ELK. Gre za zbirko orodij Logstash, Elasticsearch in Kibana. Orodje Kibana omogoča grafično predstavitev podatkov in s tem optimizacijo spremljanja kakovosti VoIP telefonije.

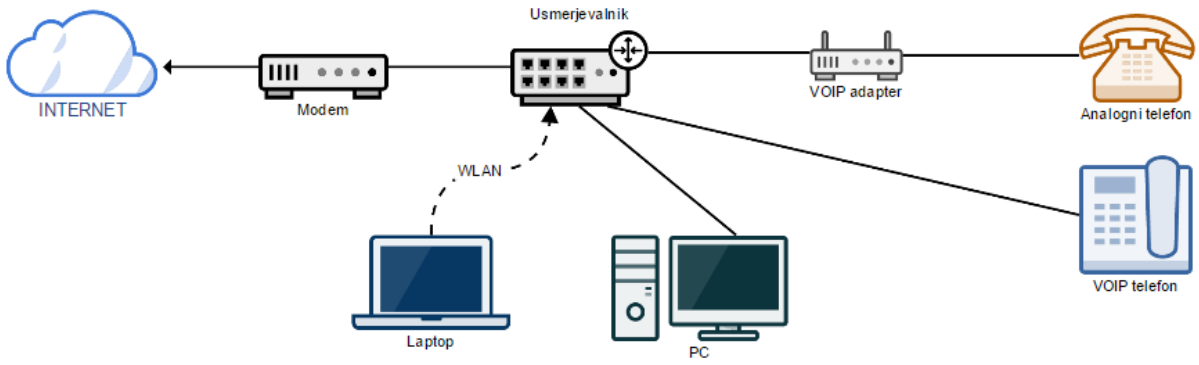
1.2 Struktura naloge

V poglavju 2 opišemo protokole internetne telefonije in parametre kakovosti glasovnih klicev. V sledečem poglavju 3 podamo raziskavo obstoječih sistemov in zahteve sistema. Nato v poglavju 4 predstavimo obstoječe tehnologije in orodja. Sledi rešitev v poglavju 5, na koncu pa sklepne ugotovitve v poglavju 6.

Poglavje 2 Internetna telefonija

Storitev VoIP je telefonija preko protokola IP oziroma internetna telefonija. Je skupek tehnologij za zagotavljanje govornih komunikacij in multimedijskih sej preko IP omrežja. Gre za komunikacijske storitve kot so glas, faks, SMS in glasovna pošta, ki se prenašajo preko interneta namesto preko javnega telefonskega omrežja PSTN.

Načela pri telefonskih klicih so podobna tradicionalni digitalni telefoniji in vključujejo signalizacijo, nastavitev kanala, digitalizacijo analognih govornih signalov ter kodiranje. Telefonska centrala je v tem primeru programska aplikacija. Končni uporabniki pa uporabljajo storitve preko odjemalca na osebnem računalniku ali telefonu in IP priključka, kot je razvidno na sliki 2.1.



Slika 2.1: Shematičen prikaz uporabe IP telefonije.

Komunikacija poteka med IP napravami preko pošiljanja medsebojnih sporočil. Analogni govorni signal se pretvori v stisnjen digitalni format. Ti digitalni podatki se razdelijo v IP pakete in se prenesejo preko IP mreže [27].

Najbolj razširjena standarda pri VoIP za signalizacijo sta SIP in H323. Slednji izginja, operaterji vse bolj prehajajo na uporabo izključno protokola SIP. Prenos avdio in video podatkov v realnem času podpira RTP.

2.1 SIP

Protokol SIP *(*ang. *Session initiation protocol)* se uporablja za signalizacijo in nadzor multimedijskih komunikacijskih sej; za vzpostavitev, spreminjanje in prekinitev sej med končnimi točkami. Večinoma gre za internetno telefonijo in prenos SMS preko IP omrežij. Seje so video in glasovni klici, prenos podatkov in SMS sporočil, video konference ter distribucija drugih multimedijskih aplikacij. Leta 2002 je bila objavljena zadnja različica RFC 3261, katerega je oblikovala IETF medmrežna delovna skupina oz. organizacija [18].

Prav tako kot VoIP je protokol SIP del aplikacijske plasti. Oblikovan je tako, da je neodvisen od transportnega sloja. Protokol lahko izvajamo na TCP, UDP ali SCTP protokolih. Temelji na besedilu in vključuje številne elemente SMTP in HTTP.

V protokolu je definiranih več omrežnih elementov, ki skrbijo za uspešno vzpostavljanje sej in izvajanje storitev. Ločimo jih na dva dela: uporabniški agent odjemalec, ki vzpostavi klic, in uporabniški agent strežnik, ki sprejme klic. Isti element je lahko hkrati strežnik in odjemalec.

Omrežni elementi komunicirajo med seboj s SIP sporočili, ki so besedilnega tipa in so podobna sporočilom HTTP. Obstajata dva tipa sporočil SIP: zahteve in odzivi. Zahteva sproži transakcijo med dvema SIP entitetama. Ključne zahteve so INVITE, BYE in REGISTER.

Koda statusa je tromestno število, ki določa odziv klica. Obstaja več vrst odzivov, ki so razporejeni v 6 skupin. Prva števka v kodi statusa določa vrsto odziva. V tabeli 2.1 so prikazani najpogostejši odzivi [18].

Vrsta odziva	**Koda statusa**	**Opis odziva**
Informativen (označuje le napredek klica)	100 180 181 182 183	poskušanje zvonenje klic je posredovan čakalna vrsta napredek seje
Uspešen	200 202	uspešna zahteva sprejeto, zveza ni vzpostavljena
Preusmeritev	300 301 302 305 380	več možnosti preseljen, URI ni več veljaven začasno preseljen uporabi proxy strežnik alternativna storitev
Napaka na strani odjemalca	400 401 402 403	neuspešna zahteva nepooblaščeno zahtevano plačilo prepovedano

	404	ni najdeno
	405	metoda ni dovoljena
	406	ni sprejemljivo
	410	konflikt
	411	neveljavna dolžina
	481	transakcija ne obstaja
	482	zaznana zanka
	484	nepopolen naslov
	485	dvoumen naslov
Napaka na strani strežnika	500	notranja napaka strežnika
	501	zahteva ni prepoznana
	503	storitev ni dosegljiva
	504	strežnik ni odziven
	505	verzija ni podprta
Splošna napaka	600	povsod zasedeno
	603	zavrnjeno
	604	neobstoječa zahteva
	606	nesprejemljivo

Tabela 2.1: Kode statusa in vrste odzivov protokola SIP [18].

2.2 RTP in RTCP

RTP *(*ang. *Real-time Transport Protocol)* je transportni protokol v realnem času. Zagotavlja funkcije transporta za aplikacije, ki prenašajo podatke v realnem času kot so avdio, video, simulacije podatkov preko omrežnih storitev. Protokol ne obravnava rezervacije sredstev in ne zagotavlja kakovosti storitve. Dopolnjuje ga RTCP protokol, ki omogoča spremljanje pošiljanja podatkov in zagotavlja minimalen nadzor. Slednji je neodvisen od transportne in omrežne plasti [26].

RTCP (ang. *RTP Control Protocol*) je sestrski protokol protokola RTP. Njegove osnovne funkcionalnosti in struktura paketov je definirana v specifikaciji RFC 3550, ki nadomešča prvotno standardizacijo iz leta 1996, RFC 1889. Zagotavlja statistične podatke in nadzor nad informacijo za pretok RTP toka. Pomaga pri dostavljanju in pakiranju multimedijskih datotek, ampak ne omogoča transporta medijskih tokov [11].

Primarna funkcija RTCP je zagotavljanje povratne informacije o kakovosti storitve s periodičnim pošiljanjem statističnih podatkov udeležencem v multimedijski seji. Zbira statistiko o medijski seji in informacije o prenešenih oktetih, številu paketov, zgubljenih paketih, RTT. Aplikacije lahko uporabijo te informacije za nadzor kakovosti storitev.

2.3 CDR

CDR (ang. *call detail record)* je razčlenjeni zapis klica, znan tudi pod imenom SMDR (ang. *Station Message Detail Recording*). Je podrobni zapis podatkov, ki ga proizvede telekomunikacijska oprema med beleženjem podatkov o komunikaciji transakciji, kot sta telefonski klic in SMS. Zapis vsebuje različne atribute kot so začetni čas, trajanje, stanje zaključka, številka vira, ciljna klicana številka, točka originacije, točka destinacije itd.

Vsebuje metapodatke, to so podatkovna polja, ki opisujejo določeno instanco telekomunikacijske transakcije, ne vsebujejo pa vsebine transakcije. Primer: zapis določenega telefonskega klica lahko vsebuje telefonski številki kličoče in klicane strani, začetni čas klica in trajanje. Ne vsebuje vsebine pogovora [4].

Definirani so atributi vsakega zapisa. Seznam osnovnih nujno potrebnih atributov [5]:

- začetni čas: točen datum in čas,
- čas vzpostavitve klica: točen čas, ko se začne obračunavati trajanje pogovora,
- čas zaključka klica: točen čas, ko se pogovor preneha,
- trajanje klica: dolžina klica v milisekundah,
- številka za obračun: telefonska številka, ki ji je klic zaračunan,
- uspešnost dostave: rezultat klica, ki nakazuje ali je bil povezan,
- nosilec storitve,
- namen,
- identifikacijska številka: enolična številka zapisa,
- status klica: ali je bil odgovorjen ali ne,
- telefonska številka klicanega naročnika: uporabnik na strani destinacije klica, B-stran,
- telefonska številka kličočega uporabnika: uporabnik na strani izvora klica, A-stran,
- dodatna številka: dodajo se številke pred telefonsko številko B-strani zaradi usmerjanja klica,
- vrsta klica: glasovni klic, SMS, faks itd.,
- koda statusa klica,
- vrsta naročnika,
- vzrok za prekinitev klica.

Zapisi CDR so pomembni za obračun časa trajanja klicev. Tako se na osnovi zapisov o trajanju klicev z računovodskimi aplikacijami ustvarja račune. Ker nam kažejo informacije o klicih, jih lahko uporabimo za zaznavanje goljufij z iskanjem vzorcev v zapisih.

Uporabljajo se tudi za analizo kakovosti klicev. Na podlagi zapisov CDR je mogoče izračunati parametre kakovosti telefonije. Ko se zazna padec kakovosti, lahko preverimo

atribute pri raziskavi problema. Že osnovni atributi nam podajo informacije o tem, kaj se je z določenim klicem dogajalo.

2.4 Parametri kakovosti glasovnih klicev

Obstajajo statistični parametri, ki merijo kakovost klicev. Kot priporočene kazalnike za nadzor večje količine klicev jih je opisala mednarodna telekomunikacijska zveza ITU [15]. V tem poglavju so opisane tudi ostale mere kakovosti, ki jih lahko uporabimo pri grafični predstavitvi kazalnikov kakovosti VoIP telefonije.

2.4.1 Osnovni parametri

NER (ang. *network effectiveness ratio*) meri zmožnost omrežja, da dostavi klice do končnega terminala. Je mera kakovosti omrežja, ki ni odvisna od vpliva uporabnikov in obnašanja končnih terminalov. Pri izračunu se štejejo kot uspešni vsi klici, kjer se zveza ne vzpostavi, vendar dobimo odgovor s končnega omrežja, kot prikazuje enačba 2.1. V splošnem NER uporabljamo za izračun učinkovitosti mednarodnih medpovezavnih poti klicev. Želimo, da je čim višji, med 95% in 100%.

$$NER = 100\frac{\boldsymbol{vzpostavljeni\ klici + zasedeni\ klici + neodgovorjeni\ klici + klici\ zavrnjeni\ s\ strani\ končnega\ terminala}}{\boldsymbol{vsi\ klici}} \quad (2.1)$$

ASR *(*ang. *answer-seizure ratio)* meri uspešnost klicev v telekomunikacijah. Podaja razmerje med številom vseh klicev, ki se vzpostavijo in številom vseh zajetih klicev, kot prikazuje enačba 2.2. Je mera, ki je neposreden pokazatelj učinkovitosti storitve, saj šteje kot uspešne samo tiste klice, ki imajo trajanje. Vsi signali za zasedeno in ostale zavrnitve klicev štejejo kot neuspešni [16].

$$ASR = 100\frac{\boldsymbol{klici\ s\ trajanjem}}{\boldsymbol{vsi\ klici}} \quad (2.2)$$

Podana je v odstotkih, mora biti med 30% in 50%. Vse kar je nad 50% nakazuje na odlično kakovost storitve, v določenih primerih je tudi pokazatelj goljufij na večjem vzorcu klicev, ki ga imenujemo *promet*. Primer uporabe: celoten vzorec vsebuje 200 klicev, od tega jih je 65 imelo trajanje. ASR se zato izračuna kot 65/200 pomnoženo s 100. Rezultat dobimo v odstotkih in znaša 32,5%.

ACD *(*ang. *average call duration)* je mera, ki izraža povprečno trajanje klica. Izračuna se tako, da se deli celotno trajanje vseh zajetih klicev s številom pogovorov oziroma številom klicev, ki so imeli trajanje. Izračun prikazuje enačba 2.3 [15].

Temelji na primerih CDR, na podlagi katerih se lahko dostopa do podatkov o trajanju vsakega klica posebej. Meri se v časovnih enotah. Nizek ACD je lahko odraz slabe kakovosti zvoka, kar povzroča, da končni uporabniki hitreje prekinjajo klice.

$$ACD = 100\frac{\sum \boldsymbol{trajanje\ vseh\ klicev}}{\boldsymbol{vsi\ klici, ki\ so\ imeli\ trajanje}} \tag{2.3}$$

Na povprečno trajanje klica lahko vpliva mnogo dejavnikov. Večinoma so neposredno povezani s kakovostjo zvoka med tem, ko je zveza že vzpostavljena. Najbolj pogosti so prezasedenost omrežja, prevare (lažni posnetki in tajnice, ki računajo trajanje), odmev, zakasnitve, trepetanje in izguba paketov, nizko zmogljivi kodeki.

2.4.2 Ostale mere kakovosti

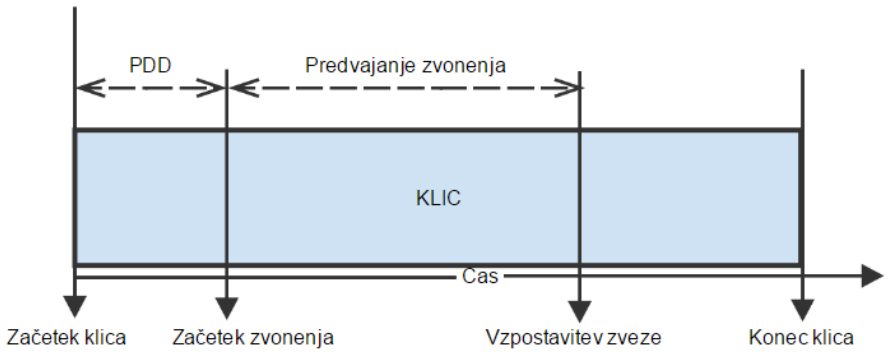
ABR (ang. *answer bid ratio*) podaja razmerje med številom klicev, ki se vzpostavijo in številom vseh klicev, ki so prišli do končnega omrežja. Lahko rečemo, da poda razmerje med ASR in NER. Izračun je prikazan v enačbi 2.4. Meri se v odstotkih in v določenih primerih, ko je NER nizek, kaže bolj realno sliko o uspešnosti klicev kot ASR [15].

Primer: Vzorec 1000 klicev, od tega jih zaradi napak na omrežju do končnega terminala pride le 400, 200 uporabnikov se oglasi. ASR je v tem primeru le 20%, med tem ko je ABR 50%. Iz ABR je razvidno, da je uspešnost povezanih klicev v mejah normale, le prepustnost omrežja je nizka. Če bi opazovali samo kazalnik ASR, bi zmotoma mislili, da je le 20% klicev bilo povezanih.

$$ABR = 100\frac{\boldsymbol{klici\ s\ trajanjem}}{\boldsymbol{vzpostavljeni\ klici + zasedeni\ klici + neodgovorjeni\ klici + klici\ zavrnjeni\ s\ strani\ končnega\ terminala}} \tag{2.4}$$

PDD *(*ang. *post dial delay)* je zakasnitev zvonenja, ki jo izkusi uporabnik na strani izvora klica (A-stran) in se meri v časovnih enotah. To je čas tišine, ki traja med vpisano zadnjo številko klicane telefonske številke in odzivom s strani uporabnika na strani ponora klica. Ta

odziv je lahko zvonenje ali pa posneto obvestilo s strani končnega operaterja. Nižji kot je PDD, boljša je uporabniška izkušnja.



Slika 2.2: Prikaz PDD.

MOS *(*ang. *mean opinion score)* je mera kakovosti zvoka od 1 do 5, kjer 1 pomeni slaba kakovost zvoka z veliko šuma ali celo tišina in nerazumljiv govor, med tem ko 5 pomeni odlična kakovost brez motenj. Lestvica MOS je navedena v tabeli 2.2.

MOS	**Pomen**
5	Odlična
4	Dobra
3	Še ustrezna
2	Slabša
1	Nesprejemljivo slaba

Tabela 2.2: Lestvica MOS

Lestvico je februarja leta 2001 standardizirala ITU v modelu P.862 za izračun kakovosti klicev z metodo zaznavnih ocen kakovosti govora PESQ (ang. *Perceptual Evaluation of Speech Quality)*. Izvedli so serijo testov, kjer so uporabniki različnih starosti in ras poslušali vrsto zvočnih posnetkov ter jih razvrščali glede na kakovost zvoka. Po tovrstnih začetnih subjektivnih rezultatih so bili razviti algoritmi, ki se sedaj uporabljajo za mero kakovosti zvoka z zornega kota uporabnika [14, 24].

Ocena MOS se pogosto uporablja za subjektivno oceno kakovosti zvoka. Metoda PESQ potrebuje meritve, ki jih pridobimo s testnimi klici, kar je razvidno iz enačbe 2.5. Poslušalec oceni vsak primer z eno izmed zgodnjih kategorij kakovosti zvoka. Gre za vsoto ocen, kjer je R individualna ocena N subjektov.

$$MOS = \frac{\sum_1^N \boldsymbol{R_i}}{\boldsymbol{N}} \tag{2.5}$$

Ocena je definirana le z besedami v tabeli. Njena slabost je ohlapna definicija MOS lestvice, ki je zato občutljiva na proceduro poslušanja. Vrednosti, ki jih pridobimo pri tem testu so odvisne od šumov v ozadju poslušalcev, kakovosti posnetkov, ki jih poslušajo, nivoja glasnosti in opreme, ki se uporablja pri poslušanju. Zato se lahko ocena MOS spreminja pri različnih testih na enakem telefonskem prenosnem sistemu.

POLQA (ang. *Perceptual Objective Listening Quality Assessment)* je novejši model za napovedovanje kakovosti zvoka, ki ga je razvila ITU [13]. Uvaja izboljšave v točnosti določanja kakovosti zvoka in kredibilnosti. Aplicira se na najnovejših kodekih za zvok, podpira testiranje pri visoki stopnji šuma v ozadju. Algoritem POLQA vsebuje dva modela NB in SWB v povezavi z različnimi primeri zvoka. Postal je priporočen algoritem leta 2011.

Primeri vrednosti MOS: telefonsko omrežje, ki ga večina ljudi uporablja doma ima povprečje MOS 4,3. GSM omrežje ima povprečje MOS med 2,9 in 4.1. Pri VoIP klicih so vrednosti MOS med 3.5 in 4.2.

Poglavje 3 Obstoječe rešitve

Podjetje Mobik se je odločilo, da sistemi za nadzor infrastrukture, kot sta *Nagios* in *Cacti,* niso primerni za grafični prikaz kazalnikov kakovosti klicev. Narejena sta za spremljanje naprav in omrežja, v podjetju pa potrebujejo analitično orodje.

Pri raziskavi obstoječih orodij smo naleteli na različne sisteme za podporo klicnim centrom in VoIP telefonijo znotraj podjetij kot sta *iVOIP centrala* in *COCOS CEP*.

iVOIP centrala podjetja Voco d.o.o. uporablja SIP protokol. Nanjo se lahko poveže uporabnik s SIP odjemalcem in internetno povezavo. Omogoča standardne funkcije, kot so posredovanje klicev, snemanje pogovora, kaskadni klici in glasovna pošta ter nekaj naprednih funkcij. Slednje omogočajo statistiko in poročila o čakalnih vrstah ter poročila o klicih, kot na primer katere številke se frekventno kličejo [17].

Platforma *COCOS CEP* podjetja CDE nove tehnologije d.o.o. vsebuje podporo za kontaktni center in upravljanje stikov s strankami. Aplikacija omogoča upravljanje s klici, usmerjanje prometa med moduli agentov kontaktnega centra, nadzor stanja čakalnih vrst in presežkov prometa [8].

Podjetja uporabljajo tovrstno telefonijo zaradi znižanja stroškov telefonskih klicev. Gre za sisteme namenjene centrali VOIP klicev, kjer imajo posamezni uporabniki internetno povezavo in VOIP odjemalec. Ne nudijo pa nadzora nad kakovostjo klicev.

Prav za analizo glasovnih klicev in nadzor parametrov kakovosti na spletu ni mogoče najti veliko orodij. Med raziskavo je bilo mogoče najti le štiri: *Carrier Cockpit* podjetja Ascade, *Captura Voice* podjetja Oculeus, *Alaris inVoice in Arptel Monit.*

Carrier Cockpit omogoča mobilnim operaterjem usmerjanje klicev, obračun, poravnavo in operativni nadzor nad kakovostjo klicev [7].

Captura Voice nudi celoten sistem urejanja cenikov, usmerjanja prometa klicev, obračuna in spremljanja klicev, nadzora za porabljen limit strank ter iskanje po zapisih klicev CDR [6].

Alaris inVoice je celosten sistem za večje organizacije, ki želijo analizo komunikacijskih storitev. Ponuja storitve za tri različne oddelke v podjetjih: za prodajo, center za nadzor omrežja ter za finance [1].

Arptel Monit je orodje za nadzor omrežja in VoIP klicev. Funkcionalnosti so pregled števcev o ključnih vrednostih in spremljanje mejnih vrednosti parametrov po časovnem intervalu. Vgrajen ima sistem avtomatskega pošiljanja poročil preko elektronske pošte [2].

Funkcionalnosti, ki jih slednja orodja omogočajo so:

- dostop do zapisov klicev CDR.
- Uporaba na Windows platformi in možnost spletne aplikacije.
- Poročila in povzetki analize izbranih klicev. Uporaba dodatka za Microsoft Excel, ki omogoča pregled statistike.
- Zmožnost urejanja pogleda zapisov klicev CDR in sortiranje podatkov pri klicih po časovnem razporedu ter lestvični diagram.
- Tabele in grafi, ki prikazujejo številko vzpostavljenih klicev, trajanje klicev, št. sporočil ipd.

Vsa ta orodja omogočajo tabelarično predstavitev podatkov, manjka pa grafična predstavitev. Peščica teh orodji za analizo klicev ponuja še dodatno možnost obveščanja in alarmov ter nastavljanje mejnih vrednosti pri katerih posvetijo alarmi, integracijo z obstoječimi sistemi za testiranje klicev, grafikone kakovosti klicev ASR in ACD, števila klicev in števila minut. Na tem mestu raziskave ni bilo mogoče najti nobenih konkretnih primerov uporabe grafikonov kazalnikov ASR, NER in ACD. Podan je bil le opis podjetja Arptel Ltd., da njihovo orodje omogoča grafikone kakovosti klicev [3].

3.1 Zahteve sistema

Po analizi obstoječih postopkov nadzora VoIP telefonije smo definirali glavne zahteve za optimizacijo nadzora nad kakovostjo klicev, ki bodo pomagale centru za nadzor prometa hitreje zaznati padec kakovosti na določeni destinaciji. Zahteve sistema so sledeče:

Tehnologija, ki ni plačljiva in jo lahko po želji oblikujemo ter nadgradimo po svojih potrebah. Mora omogočati razširljivost in združljivost z obstoječim sistemom. Eden izmed

pomembnih odločitvenih kriterijev je cena, ki mora biti čim nižja. Zato že v izhodišču iščemo odprtokodne rešitve.

Grafična predstavitev: grafi in diagrami nam pomagajo za lažje in hitrejše odkrivanje napak. Tabelarični pogled CDR že imamo, potrebujemo še vizualno predstavitev za lažje zaznavanje slabe kakovosti klicev pri operaterjih. Rešitev mora obsegati:

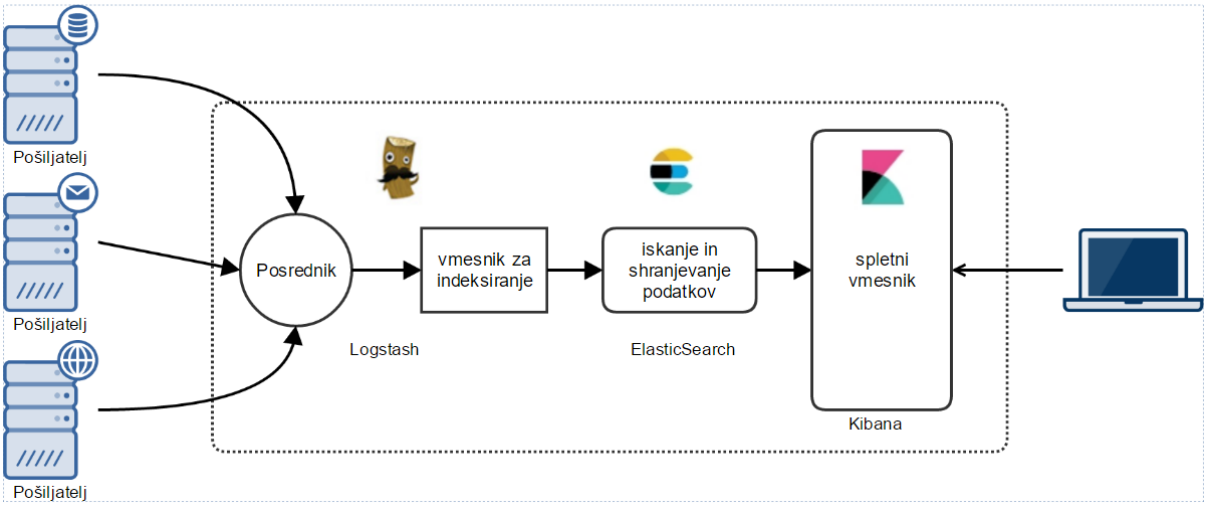
- prikaz kazalnikov kakovosti klicev na prometu končnega omrežja določenega operaterja.
- Prikaz porazdelitve kode statusa SIP klicev. Vsak klic vsebuje informacijo o odzivu, ki sporoča, če je bil klic uspešen ali pa je prišlo do napake.
- Geografski prikaz originacije klicev in. To nam v določenih primerih pomaga zaznati prevare (SPAM promet), poleg tega spremljamo strukturo oziroma profil prometa.

Spremljanje kazalnikov kakovosti po časovnem intervalu: imamo sistem, ki nam na 20 minutne časovne intervale prikazuje klice. Potrebujemo orodje, ki uspe v realnem času procesirati časovno vrsto CDR primerov in grafično prikazovat statistiko.

Poglavje 4 Uporabljene tehnologije in orodja

Dodatek k sistemu za spremljanje in analizo klicev smo implementirali z zbirko orodij Logstash, Elasticsearch in Kibana. Na sliki 4.1. je podana arhitektura sklada ELK.

V tem poglavju najprej opišemo tehnologijo virtualizacije Docker, ki nam omogoča nemoteno delovanje sklada ELK. Nato opišemo orodje Logstash, ki preoblikuje podatke, podatkovno zbirko Elasticsearch ter predstavimo Kibano, orodje za grafično predstavitev podatkov iz Elasticsearch.



Slika 4.1: Prikaz arhitekture sklada ELK [9].

4.1 Docker

Docker je odprtokodni projekt tehnologije, ki avtomatizira razmestitev aplikacij znotraj tako imenovanih programskih vsebnikov (ang. *Docker container*). Deluje po principu virtualizacije na nivoju operacijskega sistema. Kos opreme ovije v datotečni sistem, ki vsebuje vse potrebno za zagon: kodo, sistemska orodja, knjižnice. Zagotavlja, da bo programska oprema vedno enako delovala, neglede na to, v katerem okolju je nameščena.

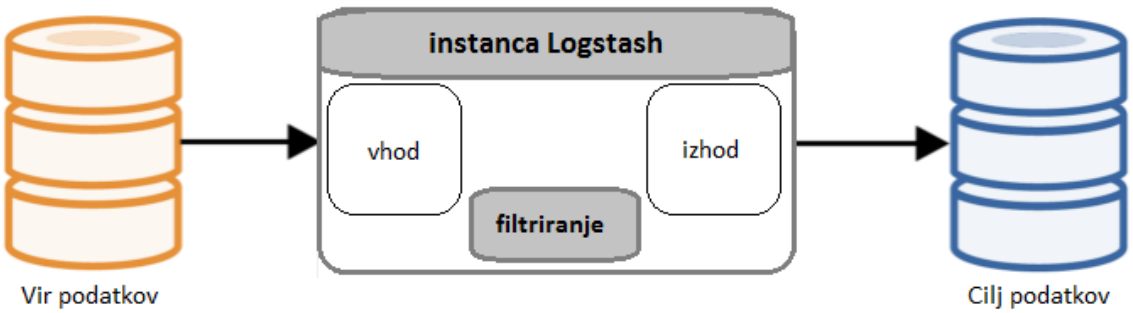
Zagotavlja dodatno plast abstrakcije in avtomatizacijo virtualizacije na operacijskih sistemih Windows ali Linux. Omogoča, da vsebniki neodvisno delujejo na eni instanci operacijskega

sistema, s tem da se izognemo zaganjanja in vzdrževanja virtualnih strojev. Uporablja različne vmesnike, da lahko dosega značilnosti virtualizacije na jedru Linux [28].

4.2 Logstash

Logstash je odprtokodni zbiralec besedilnih podatkov, namenjen je za združevanje podatkov in njihovo obdelavo. Lahko ga nastavimo tako, da dinamično poenoti podatke iz neenakih virov in jih normalizira. Torej »očisti« za nadaljnje analize ali grafični prikaz podatkov. Napisan je v programskem jeziku Ruby, lahko ga razširimo z vtičniki.

Prvotno je bil namenjen za zbiranje dnevniških zapisov, sedaj njegove zmogljivosti to presegajo. Z uporabo Logstash lahko različne tipe vhodnih podatkov po želji preoblikujemo. To storimo s pomočjo filtrov. Ko so podatki obdelani, jih pošljemo na izhod s pomočjo izhodnih vtičnikov, kot je prikazano na sliki 4.2.



Slika 4.2:Potek obdelave podatkov.

To orodje združuje podatke, obdela vhodne podatke in jih vrne na izhod. Delamo z nastavitvenimi datotekami, kjer nastavljamo nekatere parametre, spreminjamo sintakso in nastavljamo filtre [22]. Na sliki 4.3 je prikazano ogrodje nastavitvene datoteke za Logstash.

```
1  Input
2  {
3          ##34 različnih tipov vhodov
4  }
5  Filter
6  {
7          ##29 različnih filtrov
8  }
9  Output
10 {
11         ##47 različnih izhodov
12 {
```

Slika 4.3: Struktura nastavitvene datoteke Logstash.

4.2.1 Vhod in izhod

Podatki so pogosto razpršeni preko sistemov v različnih formatih. Logstash podpira različne tipe vhodnih podatkov, kot so TCP, UDP, preusmeritev iz datoteke ali standardnega vhoda. Omogoča pretakanje različnih podatkov hkrati. Lahko gre za dnevniške zapise, meritve, podatke spletnih aplikacij, baz podatkov, storitev AWS itd. Za vhodne podatke lahko uporabimo skoraj petdeset različnih tehnologij, lokacij in storitev, s katerih lahko preusmerimo dogodke. Vključujoč nadzorne sisteme kot je *collectd*, podatkovne baze kot je Redis in storitve kot je Twitter. Vhod je začetna točka vsake nastavitvene datoteke. V primeru, da ga ne definiramo, Logstash privzeto vzame standardni vhod *stdin*.

Nastavimo ga tako, da definiramo iz kje bomo vzeli podatke [22].

Omogoča neposreden ponor podatkov na Elasticsearch in druge različne izhode. Zaradi možnosti razširitve z več kot 200 vtičniki, lahko podatke vrne v različnih oblikah, kot so: TCP, dokument CSV ali XML, preusmeritev na standardni izhod, HTTP, IRC, SMTP, AMQP itd. Podatke lahko pošlje v Amazon storitev za shranjevanje, za syslog strežnik, podatkovno zbirko REDIS, lahko jih pošlje v obliki pripravljeni za Nagios.

4.2.2 Filtri

Omogočajo razčlenitev in preoblikovanje podatkov. Ko podatki potujejo od vira, Logstash filtri razčlenijo vsak dogodek, identificirajo in naredijo poimenska polja, da zgradijo strukturo. Preoblikovani podatki se stekajo v skupen format za lažjo analizo.

Logstash dinamično spreminja obliko in pripravi podatke ne glede na strukturo in kompleksnost. Filtre uporabimo odvisno od potreb, nekaj najpogostejših je navedenih v tabeli 4.1.

Naziv filtra	Delovanje
CSV	Pretvori vrednosti ločene z ločilom v individualna polja.
DATE	Poenoti podatke datuma v enoten format za nadaljnjo obravnavo.
FINGERPRINT	Omogoča spreminjanje in dodajanje vrednosti polj.
GEOIP	Pridobi geografske koordinate iz IP naslovov.
GROK	Naredi strukturo iz nestrukturiranih podatkov.
KV	Samodejno razčleni podatke tipa x=y, kjer je y neka vrednost.
METRICS	Sešteva dogodke in vrne vsoto.
MULTILINE	Sporočilo iz več vrstic združi v en dogodek.
TRANSLATE	Filter za splošno iskanje in zamenjavo vrednosti.

Tabela 4.1: Uporaba najpogostejših filtrov v Logstash [21].

4.3 Elasticsearch

Je porazdeljena podatkovna zbirka za iskanje po besedilu, analizo in shranjevanje podatkov. Napisana je v programskem jeziku Java in uporablja knjižnico Apache Lucene. Je eden izmed možnih izhodov orodja Logstash. Omogoča hitro shranjevanje, iskanje in analizo velike količine podatkov.

Elasticsearch je programska platforma za iskanje po podatkih v skoraj realnem času. To pomeni, da obstaja majhna zakasnitev od trenutka, ko se indeksira dokument, do trenutka, ko se začne iskanje. Hitre rezultate iskanja dosega, ker namesto iskanja besedila išče po indeksiranih poljih.

Funkcionalnosti so dosegljive preko REST uporabniškega vmesnika. Uporabnikom omogoča analizo s pomočjo poljubnega odjemalca, kot so brskalnik, poljuben programski jezik ali ukazne lupine.

4.3.1 Osnovni pojmi

Za boljše razumevanje je potrebno razložiti osnovne pojme, ki so bistveni za implementacijo naše rešitve.

1. **Indeks** je zbirka dokumentov s podobnimi karakteristikami, je groba oblika podatkovne baze. V množici vozlišč (strežnikov) po želji določamo indekse, toliko kolikor jih potrebujemo. Indeks je definiran z imenom, ki mora biti sestavljeno iz malih tiskanih črk. Njegovo ime se nato uporablja pri sklicevanju nanj v primerih, ko opravljamo indeksiranje, iskanje, posodabljanje in brisanje po dokumentu. Znotraj indeksa lahko definiramo več tipov. Primer: lahko imamo indeks za podatke strank, drug indeks za katalog produktov, tretji indeks za druge podatke.

2. **Tip** je objekt indeksa, je njegova logična kategorija. Njegova semantika je prepuščena izbiri uporabnika. V splošnem je tip definiran za dokumente, ki vsebujejo niz skupnih področij. Primer uporabe: vse svoje podatke lahko shranimo v en sam indeks, znotraj tega indeksa pa definiramo tipe podatkov kot so čas, koda klica, IP naslov itd.

3. **Vozlišče** *(*ang. *node)* je proces, ki teče na svojem lastnem strežniku. Je del naše množice, shranjuje podatke in sodeluje pri indeksiranju ter zmožnostih iskanja. Ob zagonu se definira naključno ime vsakega vozlišča po UUID. V primeru, da ne želimo že privzetega imena, ga lahko določimo. Ime je pomembno pri administraciji. V vsaki množici imamo lahko poljubno število vozlišč.

4. **Množica** (ang. *cluster)* je zbirka enega ali več vozlišč s skupnim imenom. Skupaj vsebujejo vse podatke katere uporabljamo za indeksiranje in iskanje. Privzeto ime množice je »elasticsearch«. Ime mora biti nastavljeno pravilno, da je vozlišče lahko del množice.

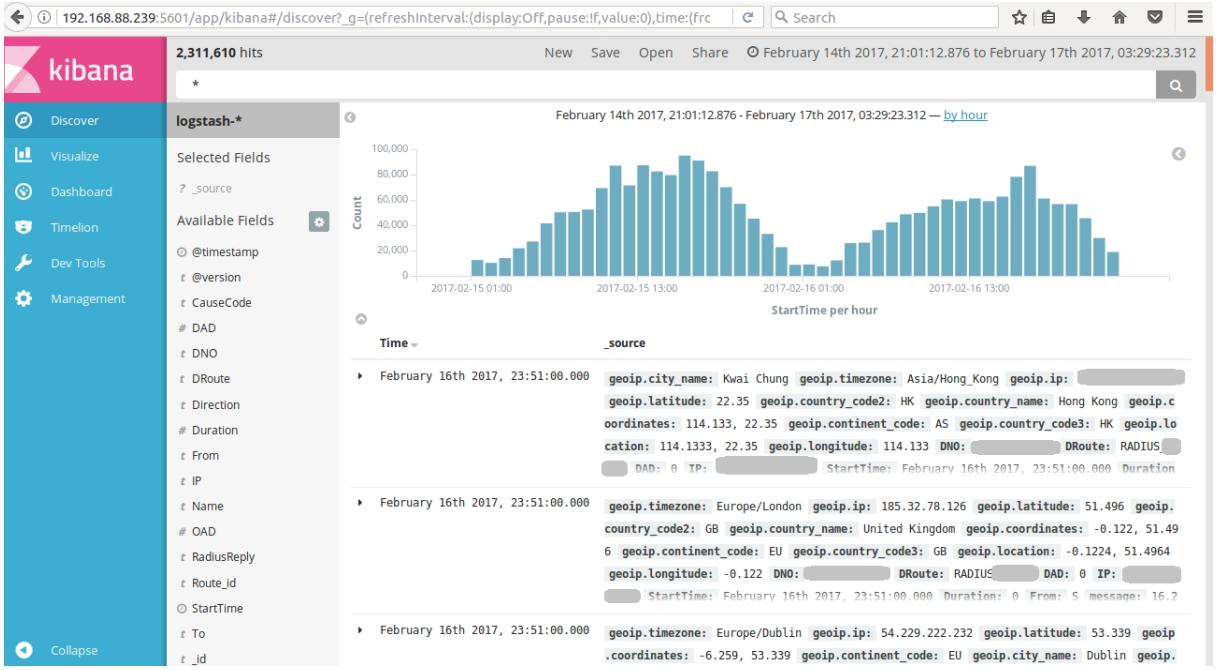
Enota iskanja in indeksiranja je dokument. Indeks vsebuje enega ali več dokumentov, dokument pa vsebuje enega ali več polj. Lahko si predstavljamo, da je dokument vrstica tabele in polje je stolpec [9].

4.4 Kibana

Kibana je odprtokodni spletni vmesnik za Elasticsearch, ki omogoča iskanje, pregled, analizo in grafični prikaz podatkov. Uporabnikom omogoča pregled po strukturi podatkov, iskanje, tvorjenje grafov in diagramov ter shranjevanje le-teh v nadzorno ploščo.

Je spletna aplikacija, dosegljiva preko vrat 5601. Za dostop do nje potrebujemo IP naslov strežnika, kjer se izvaja proces Kibane in številko vrat. Uporabniški vmesnik je razdeljen na štiri glavna področja: raziskovanje, vizualna predstavitev, nadzorna pregledna plošča in nastavitve, kot je razvidno na sliki 4.4.

Ko prvič dostopamo do orodja, se odpre področje *Raziskovanje*. Privzeto ta stran prikaže zadnje dnevniške zapise iz podatkovne zbirke Elasticsearch. V histogramu na sliki 4.4 je razvidno, kdaj smo v Elasticsearch vstavili podatke, prikazani so glede na indeks, ki ga poimenujemo. Kasneje lahko v Kibani spreminjamo indeks za različne vhodne podatke. Na tem mestu lahko z iskalnimi poizvedbami filtriramo specifične podatke. Rezultate iskanja lahko nato še bolj točno določimo s časovnim filtrom [19]. Omogoča tudi bolj zahtevno iskanje po podatkih v dodatnem področju imenovanem *Dev Tools*.



Slika 4.4: Elementi uporabniškega vmesnika Kibane in področje »raziskovanje«.

Poglavje 5 Rešitev

Celotna rešitev je realizirana s pomočjo orodja za virtualizacijo Docker, kamor smo namestili sklad ELK. Podajamo razlago arhitekture sistema z vsemi vključenimi komponentami.

5.1 Nastavitve sklada ELK

Rešitev je implementirana s pomočjo orodja Docker Compose, ki je nameščen na operacijski sistem Linux Ubuntu. Na sliki 5.1 je prikazana nastavitvena datoteka *docker-compose.yml* za sklad ELK. V njej so nastavljene uporabljene storitve: *elasticsearch, logstash* in *kibana*, ki jih definiramo v vrsticah od 3 do 35.

V razdelku *build* določimo pot za zagon storitev. V razdelku *volumes* v vrsticah od 17 do 19 definiramo pot do nastavitvene datoteke Logstash in vhodnih podatkov. Izpostavimo vrata 5000 za Logstash vhod preko TCP, 9200 in 9300 za HTTP in TCP prenos ter vrata 5601 za dostop do Kibane. V vrstici 10 nastavimo še maksimalno velikost, ki jo lahko storitev Elasticsearch zasede, da omejimo rabo bralno-pisalnega pomnilnika. To storimo preko spremenljivke okolja ES_JAVA_OPTS s parametrom Xmx.

Docker smo uporabili, ker olajša nastavitev in zagon aplikacij. Za zagon vseh treh storitev hkrati uporabimo ukaz *docker-compose up*.

```
3  services:
4    elasticsearch:
5      build: elasticsearch/
6      ports:
7        - "9200:9200"
8        - "9300:9300"
9      environment:
10       ES_JAVA_OPTS: "-Xms1g -Xmx1g"
11     networks:
12       - docker_elk
13   logstash:
14     build: logstash/
15     command: -f /etc/logstash/conf.d/
16     volumes:
17       - ./logstash/config:/etc/logstash/conf.d
18       - ./logstash/in:/logs
19       - ./logstash/var/log/:/var/log/logstash
20     ports:
21       - "5000:5000"
22     networks:
23       - docker_elk
24     depends_on:
25       - elasticsearch
26   kibana:
27     build: kibana/
28     volumes:
29       - ./kibana/config/:/etc/kibana/
30     ports:
31       - "5601:5601"
32     networks:
33       - docker_elk
34     depends_on:
35       - elasticsearch
36
```

Slika 5.1: Nastavitvena datoteka docker-compose.yml.

5.1.1 Filtriranje vhoda

Za primer diplomskega dela je vir vhodnih podatkov več različnih dokumentov. Podatke smo črpali iz strukturirane podatkovne baze z SQL poizvedbami.

Iz vrstice 3 in 4 na sliki 5.2 je razvidno, da lahko v nastavitveni datoteki orodja Logstash za vhodne podatke vzamemo enega ali več dokumentov tipa CSV. Definirano je, kje naj se začne branje in tip vhoda. Le-ta je objekt indeksa.

```
1  input {
2      file {
3          #path => "/logs/CDRprimeri.csv"
4          path => "/logs/*.csv"
5          start_position => "beginning"
6          type => "csv"
7      }
8
9  }
10
11 filter {
12     csv {
13         autogenerate_column_names => false
14         columns => ["StartTime","Name","Direction","DNO","IP","OAD","DAD","RadiusReply",
15         "Duration","DRoute","Route_id","alerting_start_time","From","CauseCode","To"]
16         separator => ";"
17         skip_empty_columns => true
18         periodic_flush => true
19     }
20
21     date {
22         match => ["StartTime", "d.M.yyyy H:mm", "yyyy-MM-dd HH:mm:ss", "ISO8601"]
23         target => "StartTime"
24     }
```

Slika 5.2: Določitev vhoda v Logstash.

Vtičnik z imenom *csv* je del filtra. Struktura polj v dokumentih je vedno enaka, zato jih po želji poimenujemo. Polja so privzeto tipa *String*, zato je potrebno ročno določiti tipe polj za kasnejšo obdelavo v Kibani. Zaradi več možnih zapisov datuma in časa, ju je potrebno definirati in pretvoriti v polje *date*, ki ima enoličen zapis. Slednje je razvidno iz vrstice 22 na sliki 5.2.

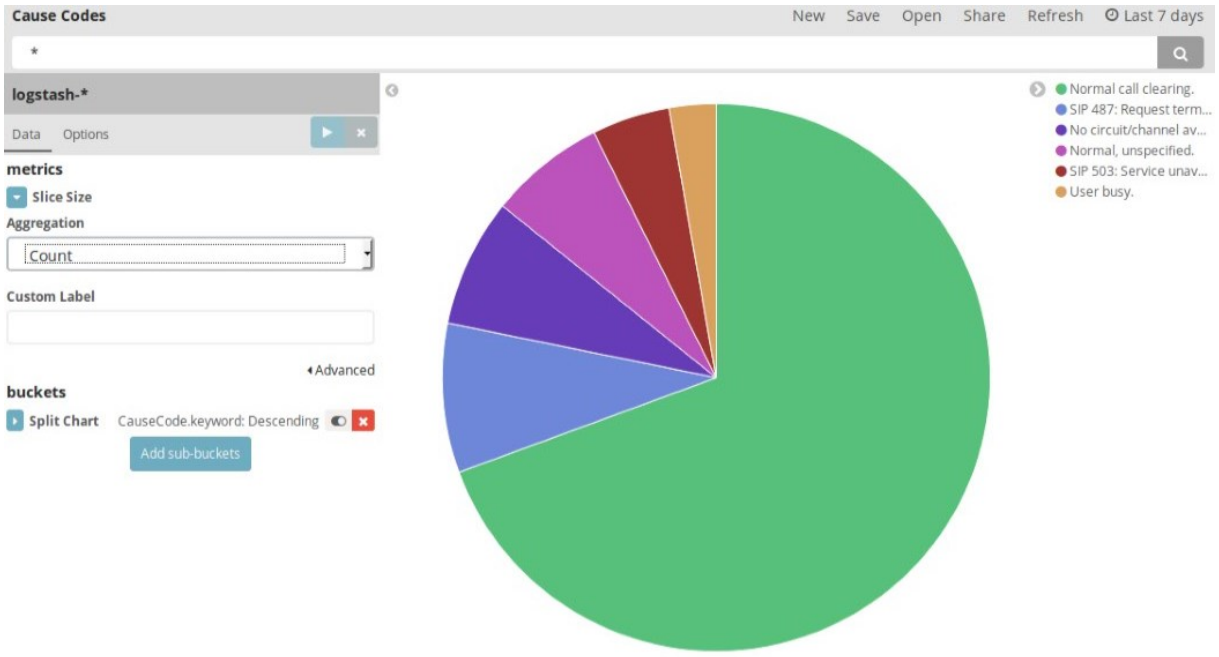
Dodatno je uporabljen filter GEOIP za upravljanje z IP naslovi klicev. V vrsticah od 73 do 78 na sliki 5.3 določimo polje *geoip* in pot do njene podatkovne baze. Uporabili smo podatkovno bazo *GeoLite2,* kjer je podprta natančnost le do nivoja mest [10].

```
67     mutate {
68         gsub => [
69             "IP", ":(.*)", ""
70         ]
71     }
72
73     geoip {
74         source => "IP"
75         target => "geoip"
76         database => "/geoipdb/GeoLite2-City.mmdb"
77         add_field => [ "[geoip][coordinates]", "%{[geoip][longitude]}" ]
78         add_field => [ "[geoip][coordinates]", "%{[geoip][latitude]}"  ]
79
80     }
81
82     #koordinatam pretvorimo tip v float
83     mutate {
84       convert => [ "[geoip][coordinates]", "float"]
85     }
```

Slika 5.3: Uporaba filtra GEOIP.

5.1.2 Grafični prikaz kazalnikov

Kot je razvidno iz iskalnega niza v brskalniku na sliki 4.4, se do Kibane dostopa preko *http://localhost:5601*. V rešitvi uporabljamo privzeti indeks z imenom *logstash*. Zaradi vgrajene podpore združevanja podatkov, je poenostavljeno ustvarjanje grafičnih predstavitev podatkov. Določene podatke smo tako filtrirali po poljih, katerim smo predhodno nastavili tip. Primer: polje *CauseCode* prikazano v tortnem diagramu, slika 5.4.

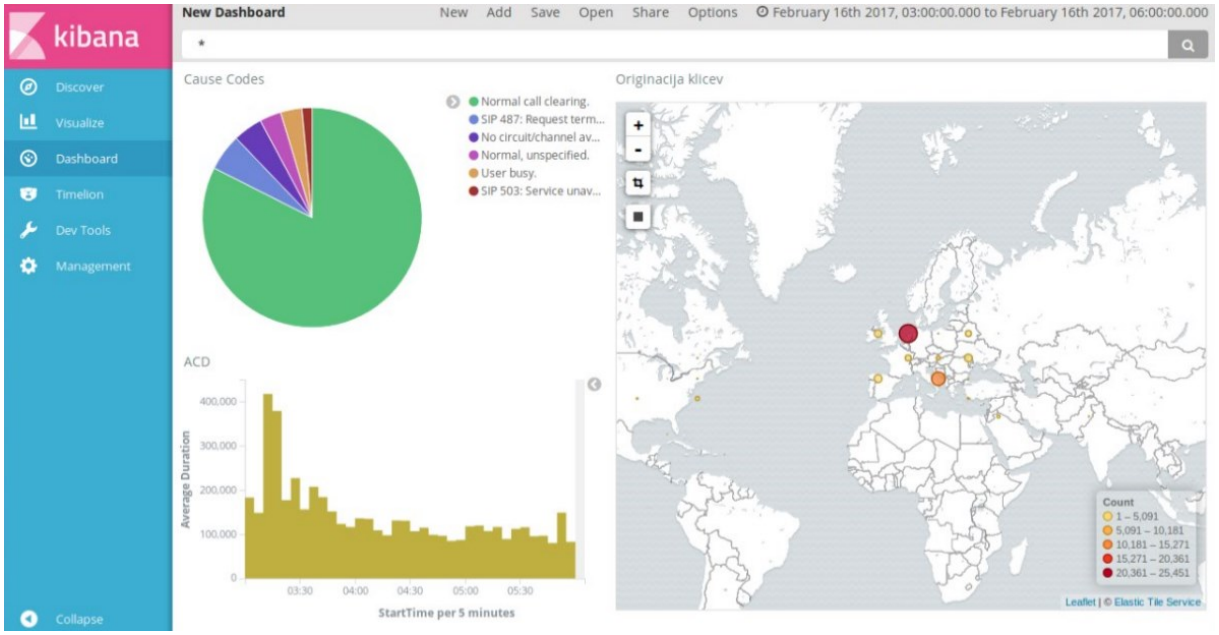


Slika 5.4: Tortni diagram za primer polja *CauseCode*.

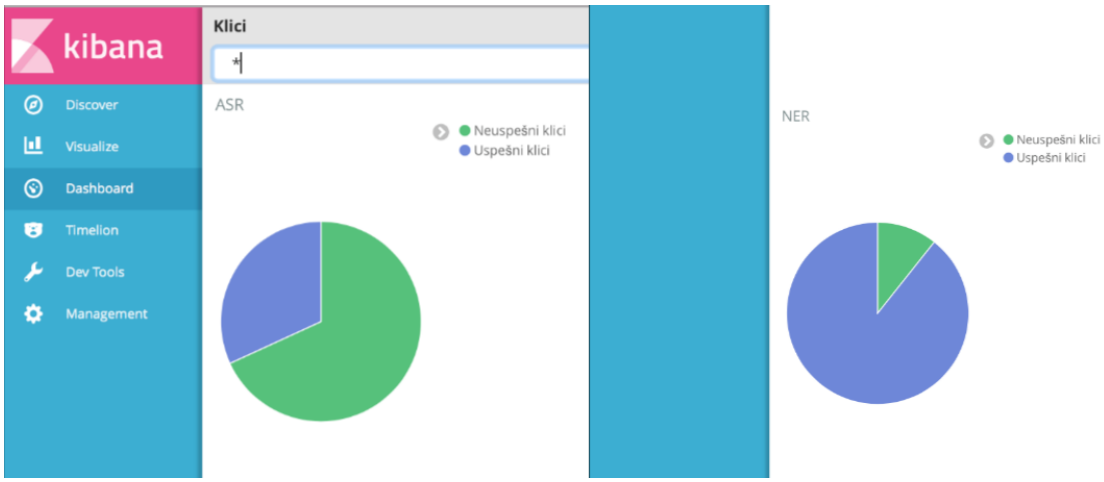
Za izris bolj kompleksnih poizvedb po bazi Elasticsearch pa je potrebno dodatno rokovati s podatki. Za upravljanje s podatki, kot je izračun kazalnikov kakovosti, smo uporabili vmesnik znotraj razdelka *Dev Tools* v Kibani. Na sliki 5.5 je podan primer, ki je napisan v JSON nastavitveni datoteki. Poizvedbe po bazi Elasticsearch smo nato uporabili za izdelavo filtrov, ki jih je mogoče uporabiti kot tip združevanja pri prikazovanju grafov.

Slika 5.5: Primer iskanja po podatkih v JSON nastavitveni datoteki.

Vse končne grafe se lahko prikaže skupaj na področju nadzorne pregledne plošče (ang. *dashboard*). Kibana omogoča shranjevanje preglednih plošč tako, da jih lahko uporabniki uporabljajo na različnih vhodnih podatkih. Na slikah 5.6 in 5.7 podajam primer končnih podob nadzorne plošče.



Slika 5.6: Prvi primer shranjene pregledne plošče v Kibani.



Slika 5.7: Drugi primer shranjene pregledne plošče v Kibani.

5.2 Ovrednotenje rešitve

V rešitvi smo prikazali funkcionalnosti sklada ELK uporabljene v diplomskem delu. Sestavni del so Docker, Docker Compose, nastavitve Logstash, indeksiranje in grafični prikaz v Kibani ter JSON nastavitvene datoteke za poizvedbe v Elasticsearch. Predlagana rešitev zadošča zahtevam sistema, ki so opisane v poglavju 3.

Tehnologija, ki ni plačljiva: uporabljena orodja in tehnologije so brezplačne. Dovolj je manevrskega prostora za nadgradnjo in integracijo z obstoječim sistemom. Na tem mestu je potrebno izpostaviti, da sklad ELK omogoča pošiljanje izhodnih datotek neposredno na Zabbix strežnik namesto v orodje Kibana [23]. *Zabbix* je odprtokodni program za nadzor omrežij in aplikacij. Narejen je tako, da spremlja status spletnih storitev, strežnikov in ostale programske ter strojne opreme na omrežju. Uporablja MySQL, PostrgesSQL, Oracle in IBM DB2 za shranjevanje podatkov, njegov spletni vmesnik pa je napisan v PHP.

Mogoča bi bila nadgradnja rešitve za integracijo z Zabbix strežnikom, vendar bi to za nov nadzorni sistem kakovosti klicev pomenilo, da je potrebno uporabiti uporabniški vmesnik orodja Zabbix. Le-ta ne omogoča toliko funkcionalnosti kot Kibana, ki je narejena izključno samo za grafični prikaz podatkov. Kot je zapisano v dokumentaciji, Zabbix omogoča prikaz vgrajenih preprostih grafov [29]. Uporabiti je možno le črtni in ploščinski grafikon, ki sta namenjena za prikaz podatkov, kjer na abscisni osi spremljamo čas. Kibana omogoča izdelavo in prikaz tako preprostih metrik in tabel podatkov, kot tudi izris podatkov v črtnem, stolpčnem, tortnem, kolobarnem in paličnem grafikonu, toplotnem grafu, ter na zemljevidu.

Grafična predstavitev: do sedaj je v sistemu obstajala le tabelarična predstavitev podatkov, grafično predstavitev smo dosegla z orodjem Kibana.

1. Prikaz **porazdelitve kode statusa SIP klicev** je prikazan na sliki 5.4. V primeru, da smo primerno strukturirali podatke v nastavitveni datoteki Logstash, jih je v Kibani mogoče enostavno prikazati s pomočjo filtra po izbranem polju.

2. **Geografski prikaz** originacije klicev za zaznavanje prevar je prikazan na sliki 5.6. Poda strukturo oziroma profil prometa klicev in pomaga centru za nadzor prometa.

3. Pri zahtevi za **prikaz kazalnikov kakovosti** VoIP telefonije smo naleteli na problem, saj Logstash ne omogoča izvajanja operacij nad podatki in statistične obdelave. Je orodje za rokovanje z vhodnimi besedilnimi podatki za nadaljnji prikaz v Kibani ali za izpis na drugo vrsto izhoda. Za izračun kazalnikov smo poizkušali v dokumentih ustvariti nova polja z vključevanjem dodatnih modulov znotraj filtrov. V ta polja smo želeli shranjevati

dodatno informacijo o uspešnosti klicev, ki bi se izračunala glede na vsebino ostalih polj. Primer: Modul *ruby* ima opcijo *add_field*, s katero je omogočeno dodajanje polj. Slednje se lahko izračuna s programskim jezikom Ruby. Izkazalo se je, da je to preprostejše storiti z urejanjem JSON datotek, kar je prikazano na sliki 5.5.

Spremljanje kazalnikov kakovosti po časovnem intervalu: Kibana omogoča učinkovit način prikaza informacij o izvoru/ponoru klicev ter kazalnikov kakovosti VoIP. Sklad ELK omogoča prikaz kazalnikov v skoraj realnem času. Do krajšega zamika prihaja zaradi analize v bazi podatkov Elasticsearch. V rešitvi je prikazano, kako se grafično prikaže podatke, za resnejšo analizo bi potrebovali dostop do strežnika in strukturirane baze za sprotno obdelavo in analizo podatkov.

Poglavje 6 Zaključek

ELK sklad je vse bolj razširjena rešitev zaradi svoje odprtosti in prilagodljivosti. V diplomskem delu smo uspešno implementirali možno rešitev za grafični prikaz kazalnikov in boljši nadzor nad kakovostjo storitve. Na primerih smo prikazali uporabo filtrov, kompleksnejše iskanje po podatkih ter končno grafično podobo. Bolj podrobno smo predstavili vse tehnologije in uporabljena orodja.

Vedoč, da se delo s skladom ELK spreminja zaradi tendence konstantnih objav novih verzij, se bo način urejanja grafov z vmesnikom *Dev tools* verjetno še spreminjal. Trenutno je najlažje podati zahtevnejše statistične izračune podatkov kar v JSON nastavitvenih datotekah za dostop do Elasticsearch.

ELK sklad je bil primarno namenjen za analizo dnevniških zapisov, zato na področju prikazovanja razčlenjenih zapisov CDR ni mogoče najti veliko literature. Naš pristop je ena izmed možnih rešitev za grafični prikaz kazalnikov kakovosti VoIP telefonije. Pripomore k hitrejšemu in lažjemu obvladovanju informacij o kakovosti klicev na večjih vzorcih.

V poglavju 5 je nakazana možnost razširitve te rešitve za združljivost s sistemom Zabbix. Vendar pa bi bilo glede na zahteve podjetja potrebno ovrednotiti, ali je smiselno rešitev uporabljati kot dodatek k orodju Zabbix. Upoštevati je potrebno možnost, da se rešitev uporablja kot ločen sistem.

Literatura

[1] Alaris inVoice. Alaris Labs. Dostopno na: http://www.alarislabs.com/company/news/18-solutions-en/contacts, 2008-2017. Zadnji dostop 27.2.2017.

[2] ArptelMonit. Arptel Ltd. Dostopno na: http://arptel.com/products_tm_arptel_call.html, 2010. Zadnji dostop 26.2.2017.

[3] Arptel Monit. Arptel Ltd. Dostopno na: http://arptel.com/products_tm_nm_ arptel_monit.html, 2010. Zadnji dostop 12.02.2017.

[4] N. Brownlee, A.Blout. Accounting Atributes and Record Formats. RFC 2924, IETF, september 2000.

[5] N. Brownlee. Specification of TMN applications at the Q3 interface: Call detail recording. Q.825, ITU-T Recommendation, 1998.

[6] Captura Voice. Oculeus GmbH. Dostopno na: http://www.oculeus.com/system capturavoice.html, 2017. Zadnji dostop 22.1.2017.

[7] Carrier Cockpit. Entico Corporation. Dostopno na: http://www.intercomms.net/FEB04/content/ascade.php, 2004. Zadnji dostop 19.1.2017.

[8] COCOS Customer Engagement Platforma (CEP). CDE nove tehnologije d.o.o. Dostopno na: https://www.cde.si/si/produkti/cocos-cep/cc-engagement-platform-3/, 2015. Zadnji dostop 21.1.2017.

[9] Elasticsearch Reference [5.2]: Basic Concepts. Elastic. Dostopno na: https://www.elastic.co/guide/en/elasticsearch/reference/current/_basic_concepts.html, 2017. Zadnji dostop 12.2.2017.

[10] GeoLite2 Free Databases. MaxMind Inc. Dostopno na http://dev.maxmind.com/geoip/geoip2/geolite2/, 2017. Zadnji dostop 13.2.2017.

[11] Glossary. Voipmonitor. Dostopno na https://www.voipmonitor.org/doc/Glossary, 2017. Zadnji dostop 15.01.2017.

[12] D. Grah. ELK stack for Hackers. Viris, varnost in razvoj informacijskih sistemov, d. o. o. Dostopno na: https://www.viris.si/2016/01/elk-stack-for-hackers, 2016. Zadnji dostop 19.2.2017.

[13] Implementer´s Guide on discrimination of wideband and superwideband speech. P.863, ITU-T, januar 2016.

[14] ITU recomendation for calculating telephone call quality using the Perceptual evaluation of speech quality (PESQ) method. P.862, ITU, februar 2001.

[15] Internal automatic observations. E.425, ITU-T, marec 2002.

[16] International network management - Operational guidance. E.411, ITU, 2000.

[17] iVOIP Centrala. Voco d.o.o. Dostopno na: http://www.ip-telefonija.net/iVOIP CENTRALA.aspx, 2007. Zadnji dostop 20.1.2017.

[18] A. Johnston, J. Rosenberg, H. Schulzrinne, G. Camarillo, J. Peterson, R. Sparks, M. Handley in E. Schooler. SIP: Session Initiation Protocol. RFC 3261, IETF, junij 2002.

[19] Kibana User Guide. Elastic. Dostopno na: https://www.elastic.co/guide/en/kibana/current/tutorial-discovering.html, 2017. Zadnji dostop 17.2.2017.

[20] L. Lhotka. JSON Encoding of Data Modeled with YANG. RFC 7951, IETF, avgust 2016.

[21] Logstash Reference [5.2]: Filter plugins. Elastic. Dostopno na: https://www.elastic.co/guide/en/logstash/current/filter-plugins.html, 2017. Zadnji dostop 15.2.2017.

[22] Logstash Reference [5.2]: Logstash Introduction. Elastic. Dostopno na: https://www.elastic.co/guide/en/logstash/current/introduction.html, 2017. Zadnji dostop 15.2.2017.

[23] Logstash Reference [5.2]: Output plugins: Zabbix. Elastic. Dostopno na: https://www.elastic.co/guide/en/logstash/current/plugins-outputs-zabbix.html, 2017. Zadnji dostop 19.2.2017.

[24] Objective quality measurement of telephone-band (300-3400 Hz) speech codecs. P.861, ITU-T, avgust 1996.

[25] Predstavitev družbe Mobik. Mobik d.o.o. Dostopno na: https://svet.fri.uni-lj.si/wp-content/uploads/2016/10/Predstavitev-dru%C5%BEbe-Mobik-d.o.o..pdf, oktober 2016. Zadnji dostop 10.2.2017.

[26] H. Schulzrinne, S. Casner, R. Frederick in V. Jacobson. RTP: A Transport Protocol for Real-Time Applications. RFC 3550, IETF, julij 2003.

[27] A. Uzelac, Y. Lee. Voice over IP (VoIP) SIP Peering Cases. RFC 6405, IETF, november 2011.

[28] What is Docker. Docker Inc. Dostopno na https://www.docker.com/what-docker, 2016. Zadnji dostop 19.2.2017.

[29] Zabbix Documentation 3.0: Graphs. Dostopno na https://www.zabbix.com/documentation/3.0/manual/config/visualisation/graphs, 2017. Zadnji dostop 26.2.2017.